DOI: 10.1016/j.peptides.2016.03.005

Structure activity relationship modelling of milk protein-derived peptides with dipeptidyl peptidase IV (DPP-IV) inhibitory activity

Alice B. Nongonierma & Richard J. FitzGerald.

Department of Life Sciences and Food for Health Ireland (FHI), University of Limerick, Limerick, Ireland).

Please cite as:

Nongonierma, A. B., & FitzGerald, R. J. (2016). Structure activity relationship modelling of milk protein-derived peptides with dipeptidyl peptidase IV (DPP-IV) inhibitory activity. Peptides, 79,1-7.

*Corresponding author: dick.fitzgerald@ul.ie
Tel: +353 (0) 61 202598
Fax: + 353 (0) 61 331490

Abstract

Quantitative structure activity type models were developed in an attempt to predict the key features of peptide sequences having dipeptidyl peptidase IV (DPP-IV) inhibitory activity. The models were then employed to help predict the potential of peptides, which are currently reported in the literature to be present in the intestinal tract of humans following milk/dairy product ingestion, to act as inhibitors of DPP-IV. Two models (z- and v-scale) for short (2-5 amino acid residues) bovine milk peptides, behaving as competitive inhibitors of DPP-IV, were developed. The z- and the v-scale models ($p < 0.05$, R^2 of 0.829 and 0.815, respectively) were then applied to 56 milk protein-derived peptides previously reported in the literature to be found in the intestinal tract of humans which possessed a structural feature of DPP-IV inhibitory peptides (P at the N_2 position). Ten of these peptides were synthetized and tested for their *in vitro* DPP-IV inhibitory properties. There was no agreement between the predicted and experimentally determined DPP-IV half maximal inhibitory concentrations (IC_{50}). However, the ranking for DPP-IV inhibitory potency of the competitive peptide inhibitors was conserved. Furthermore, potent *in vitro* DPP-IV inhibitory activity was observed with two peptides, LPVPQ (IC_{50} = 43.8 ± 8.8 µM) and IPM (IC_{50} = 69.5 ± 8.7 µM). Peptides present within the gastrointestinal tract of human may have promise for the development of natural DPP-IV inhibitors for the management of serum glucose.

Keywords. Dipeptidyl peptidase IV inhibition; structure activity modelling; bioactive peptides; milk proteins.

1 Introduction

Milk proteins contain specific peptide motifs, termed bioactive peptides (BAPs), which have been identified *in vitro* for their potential role in beneficially modulating specific biomarkers of health. To date, several dietary intervention studies have been conducted to investigate the effect of milk protein-derived peptides in human health [3, 7, 16]. While some studies have demonstrated a positive role of milk protein-derived BAPs in humans, other studies have failed to identify an improvement in specific health biomarkers following the consumption of BAPs or intact milk proteins [31, 32].

The rate of discovery of peptide sequences originating from dietary proteins which display *in vitro* bioactive properties has increased over the past number of years. This has been made possible notably with advances in chromatographic separation techniques coupled with mass spectrometric detection of peptides [36]. In addition, quantitative structure activity relationship (QSAR) studies have allowed the development of a more in-depth understanding of BAP sequences [11, 13, 34, 39]. QSAR generally consists in the analysis of multivariate peptide descriptors which are linked to biological activity using computational methods such as partial least square regression (PLSR), multiple linear regression (MLR), principal component analysis (PCA) or artificial neural networks (ANN) [10, 11]. QSAR has been applied in the study of angiotensin converting enzyme (ACE) inhibitory [14, 38, 52, 53], antimicrobial [5, 24, 40, 43], antioxidant [19, 20, 45] and bitter [15, 19, 37, 51] peptides.

The main advantage of using QSAR approaches over qualitative assessment of peptides to study their bioactive properties is that QSAR may provide a means to predict the potency of specific peptide sequences. While several bioactive properties have been studied using QSAR, others do not appear, to date, to have been assessed with this approach. As already indicated, QSAR has been applied to different BAPs, however, to date, it does not appear to have been applied to dipeptidyl peptidase IV (DPP-IV) inhibitory sequences. DPP-IV is an amino dipeptidyl peptidase which hydrolyses and inactivates the incretin hormones glucose dependent insulinotropic polypeptide (GIP), glucagon-like peptide-1 (GLP-1) and peptide YY (PYY). Inhibition of DPP-IV can prevent incretin degradation, allowing maintenance of their insulinotropic activity during the post-prandial phase [6]. While DPP-IV inhibitory drugs (i.e., gliptins) are currently used for the treatment of type 2 diabetes (T2D), there is also a growing interest in employing naturally derived compounds such as dietary peptides for their DPP-IV inhibitory properties [12, 33]. In this

context, several studies have shown that milk protein-derived peptides can inhibit DPP-IV *in vitro* [17, 25, 42, 47].

The antidiabetic properties of intact and hydrolyzed milk proteins has been reported in the literature [31, 32]. Numerous milk protein-derived peptides have been identified in the intestinal tract of humans following the ingestion of milk and dairy products [1, 2, 34]. However, the contribution of these peptides to human health is still unknown [32]. *In vivo* studies demonstrating the antidiabetic effect of milk protein-derived DPP-IV inhibitory peptides have been conducted in small animals [48, 49]. To our knowledge, no studies to date have reported that milk protein-derived peptides present within the intestinal tract of humans could act as DPP-IV inhibitors. Therefore, the aim of this study was to demonstrate that peptides which may be released during the digestion of milk proteins in humans could inhibit DPP-IV. This was achieved by compiling a database of previously identified milk protein-derived peptides in the intestinal tract of humans following the ingestion of milk and dairy products. To facilitate the evaluation of these peptide sequences, a selection of 10 synthetic peptides was assessed for their *in vitro* DPP-IV inhibitory properties within this study. Therefore, a combination of knowledge of the structural features of DPP-IV inhibitory peptides (i.e., P at the N_2 position) and a QSAR approach was used to predict the DPP-IV inhibitory potential of > 250 peptides identified in the intestine of humans.

2 Materials and methods

2.1 Reagents

Tris(hydroxymethyl)aminomethane (TRIS), trifluoroacetic acid (TFA), GP-p-nitroanilide (pNA), diprotin A (IPI) and porcine DPP-IV (≥ 10 units mg^{-1} protein) were obtained from Sigma Aldrich (Dublin, Ireland). Hydrochloric acid (HCl) was from VWR (Dublin, Ireland). The synthetic peptides (purity ≥ 95 % (w/w)) employed to build the QSAR model, i.e., FP, HP, RP, VP, IPM, LPP, IPPL and IPSK, were obtained from Thermo Fisher Scientific (Ulm, Germany) while VPGEIVE, YPFPGP, LPQNIPPLT, IPPLTQT, TPVVVPP, YPVEPF, LPLPLL, QPHQPLPPT, QPLPPT, LPVPQ were from Genscript (Piscataway, NJ, USA).

2.2 DPP-IV inhibition assay and mode of inhibition

The DPP-IV inhibition assay was conducted as described by Nongonierma and FitzGerald [25]. The peptides were tested in the concentration range of 12.5×10^{-3} to 1.25 mg mL^{-1} (final concentration). Each sample was analyzed in triplicate (n=3). The DPP-IV half maximal inhibitory concentrations (IC_{50}) were determined by plotting the percentage inhibition as a function of the concentration of test compound using a logistic regression.

The mode of DPP-IV inhibition was assessed using the Lineweaver and Burk double reciprocal representation as previously described by Nongonierma and FitzGerald [25]. The peptides were evaluated at two concentrations, i.e., their IC_{50} values divided by 8 and 40. The concentrations of GP-pNA ranged from 0.1 to 0.5 mM (final concentration). These substrate concentrations were selected as they were consistent with the limits of detection for pNA and allowed the obtention of a linear response for the Lineweaver and Burk double reciprocal analysis

2.3 QSAR of DPP-IV inhibitory peptides

The QSAR model was developed as a tool to predict the DPP-IV inhibitory potency of unknown peptides, i.e., peptides specifically reported to be present in the intestinal tract following ingestion of milk/dairy products. All IC_{50} values were obtained with the same assay as previously described for the determination of DPP-IV inhibition [25]. The peptides used to build the QSAR model consisted of competitive DPP-IV inhibitors which had been identified in previous studies (Table 1). Eight additional peptides present within milk proteins were tested for their *in vitro* DPP-IV inhibitory potential and included in the training set. Five of which (FP, HP, RP, VP and LPP) were predicted *in silico* to be of relevance to DPP-IV inhibition [26]. The z-scale (z_1 (hydrophilicity), z_2 (size) and z_3 (charge)) from Hellberg, Sjoestroem, Skagerberg and Wold [9] and the structural v-scale (v_1 (van der Waals volume), v_2 (net charge index) and v_3 (hydrophobic parameter of side chains)) developed by Lin, Long, Bo, Wang and Wu [21] were used for the amino acid descriptors. The short peptide sequences incorporated in the QSAR model ranged from 2 to 5 amino acid residues. Therefore, the peptide descriptors were generated using a method allowing data treatment for peptides with various lengths as described by Li and Li [19]. The two N- and C-terminal amino acids of the peptides were used to generate the peptide descriptors. A PLSR was used to link the DPP-IV IC_{50} values to the peptide descriptors as outlined in Equation 1:

$$Y = c + \sum_1^i \sum_1^j \alpha_{i,j} N_{i,j} + \sum_1^i \sum_1^j \beta_{i,j} C_{i,j} + \varepsilon \quad (1)$$

With Y, the DPP-IV IC_{50} value; i: the N- or C- terminal position of the amino acid (1 or 2); j: the amino acid descriptor number (varying between 1 and 3); α and β: the coefficients of the model; c, and ε the constant and residual of the model, respectively.

Generation of the peptide descriptors and the PLSR were carried out using Matlab (version R2014b, MathWorks, Inc, Natick, MA, USA). The training set for the model included 21 peptides, while 5 peptides (test set) were randomly chosen (100 times) by the Matlab algorithm for cross-validation.

2.4 Utilization of the QSAR model to predict the DPP-IV inhibitory potency of milk protein-derived peptides previously identified in the intestinal tract of humans

A database consisting of > 250 peptides reported to be released during the digestion of milk proteins in the intestinal tract of humans [1, 2, 4, 18, 22, 44] was compiled. Five peptides (YPFPGPIPN, FPGPIPN, LPQNIPPL, IPPLTQTPV and LPLPL) identified in the intestinal tract of humans have previously been reported as DPP-IV inhibitors *in vitro* (Supplementary Table S1). Out of the > 250 peptides, 56 possessed the DPP-IV inhibitory feature consisting of a P at position N_2. Predicted DPP-IV IC_{50} values of these 56 peptides were calculated using the QSAR models (Equation 1). From these 56 peptides, 10 sequences (VPGEIVE, YPFPGP, LPQNIPPLT, IPPLTQT, TPVVVPP, YPVEPF, LPLPLL, QPHQPLPPT, QPLPPT, LPVPQ) having different predicted DPP-IV IC_{50} values and sharing common sequences with previously identified DPP-IV inhibitory peptides [28] were selected for further *in vitro* studies. At the time of the study, these 10 sequences had not previously been reported for their DPP-IV inhibitory activity. The ten peptides were synthetized and evaluated *in vitro* for their DPP-IV inhibitory properties.

2.5 Statistical analysis

The statistical analysis consisted of means comparison carried out with a one way ANOVA followed by a Student Newman-Keuls test using SPSS (version 22, SPSS Inc., Chicago, IL, USA) at a significance level $p < 0.05$. For the QSAR models, the R^2 and R^2 cross validation were used to assess the significance of the PLSR.

3 Results

3.1 DPP-IV inhibitory properties and mode of DPP-IV inhibition of peptides included in the dataset

The eight additional synthetic milk protein-derived peptides (FP, HP, RP, VP, IPM, LPP, IPPL and IPSK) included in the training set were assessed for their DPP-IV inhibitory properties and their mode of inhibition. Their IC_{50} values ranged between 69.5 ± 8.7 and 902.8 ± 93.8 µM for IPM and HP, respectively (Table 1). Of these peptides, one (IPPL, Figure 1A) was uncompetitive, two (FP and HP, Figure 1B) were mixed-type, one (LPP, Figure 1C) was non-competitive and four (IPM, VP, IPSK and RP, Figure 1D) were competitive inhibitors of DPP-IV.

When peptides not behaving as competitive DPP-IV inhibitors as well as other peptide DPP-IV IC_{50} data from the literature obtained using different assay conditions were included in the training set, no significant relationship ($p > 0.05$) between the peptide descriptors and their DPP-IV IC_{50} values was obtained (data not shown). The DPP-IV IC_{50} values of the peptides included in the dataset were therefore those obtained from our previous studies [25-27, 29] with the exception RP, VP, IPM and IPSK, whose DPP-IV IC_{50} values were determined in this study (Table 1). All the DPP-IV IC_{50} values used to build the QSAR model were obtained using the same experimental conditions [25]. Furthermore, the peptides included in the QSAR dataset only consisted of the competitive DPP-IV inhibitory peptides (26) as outlined in Table 1.

3.2 QSAR models

The two QSAR models linking the DPP-IV IC_{50} value to the peptide descriptors were significant, having a p of 0.00279 and 0.00437 for the z- and v-scale models, respectively (Table 2). The R^2 and R^2 cross validation values were > 0.75 for both models, which was indicative of a relatively good correlation between the DPP-IV IC_{50} values and the peptide descriptors (Figure 2A and 2B).

The parameters of the PLSR for the two QSAR models developed with the two scales for amino acid descriptors are shown in Table 3. For the QSAR model developed with the z-scale, the two coefficients linked with the hydrophilicity of the N_1 ($\alpha_{1,1}$) and N_2 ($\alpha_{2,1}$) terminal amino acids were significant ($p < 0.05$). The positive value of these coefficients suggested that a reduced hydrophilicity for the two N terminal amino acids would decrease the DPP-IV IC_{50} value (i.e., in-

crease the potency). In the case of the v-scale model, only the coefficient linked with the hydrophobicity of the N_1 ($\alpha_{1,3}$) terminal amino acid was significant ($p < 0.05$). It was negatively correlated to the DPP-IV IC_{50} value, suggesting that hydrophobic amino acids located at the N-terminus of the peptides would tend to decrease the DPP-IV IC_{50} value. While both QSAR models resulted in different outcomes, they both predicted that the hydrophobicity of the N-terminal amino acid plays a significant role in DPP-IV inhibitory potency.

3.3 Confirmatory studies with peptides relevant to human health

Within the milk peptides identified in the intestine of humans following the ingestion of milk proteins and dairy products, 56 possessed a feature characteristic to that of previously identified DPP-IV inhibitors (i.e., P at position 2 of the peptide) as outlined elsewhere [8, 28, 46]. These peptides originated from β- and α_{s1}-casein ([1, 4], Supplementary Table S1). Within these peptides, 5 had already been reported in the literature to display *in vitro* DPP-IV inhibitory properties [29, 49], having DPP-IV IC_{50} values ranging from 46 and 1300 µM for LPQNIPPL and IPPLTQTPV, respectively (Supplementary Table S1).

The two QSAR models (Table 3) were used to predict the DPP-IV IC_{50} values of the 56 peptides (Table 4 and Supplementary Table S1). Ten (VPGEIVE, YPFPGP, LPQNIPPLT, IPPLTQT, TPVVVPP, YPVEPF, LPLPLL, QPHQPLPPT, QPLPPT, LPVPQ) of the peptides identified in humans [1] were selected for subsequent confirmatory studies. These peptides were selected on the basis that they had a P in position N_2 and contained sequences similar to previously identified DPP-IV inhibitory peptides [28]. Their DPP-IV IC_{50} values, as predicted by the 2 QSAR models, together with their DPP-IV IC_{50} values determined *in vitro* along with their mode of inhibition are shown in Table 4. All ten peptides were able to inhibit DPP-IV *in vitro*, having IC_{50} values ranging from 43.8 ± 8.8 to 1754.8 ± 375.0 µM for LPVPQ and QPHQPLPPT, respectively. Five (VPGEIVE, LPQNIPPLT, LPLPLL, QPLPPT and LPVPQ) of the peptides evaluated were competitive DPP-IV inhibitors (Table 4) while the other peptides were uncompetitive (IPPLTQT) or mixed-type inhibitors (YPFPGP, TPVVVPP, YPVEPF and QPHQPLPPT).

The z-scale QSAR model had the tendency to underestimate the IC_{50} value of the competitive DPP-IV inhibitors, while the opposite was seen with the v-scale model. Generally, there was no agreement with the DPP-IV IC_{50} values predicted with the two QSAR models and the experimentally determined IC_{50} values. This was seen in particular when negative IC_{50} values were

predicted by the z-scale model. While it is not possible to obtain negative IC_{50} values, these values were indicative of a low predicted IC_{50}. In general, the ranking of the competitive peptides in terms of their DPP-IV inhibitory potency was in agreement in both QSAR models with the experimental values (Table 4).

4 Discussion

Within this study, QSAR was used as a tool to rank peptides released in the intestinal tract of humans in relation to their ability to inhibit DPP-IV. Interestingly, statistically significant models were only obtained when incorporating data for competitive DPP-IV inhibitory peptides and for inhibitory data obtained using similar experimental conditions. This may be linked to the fact that the conditions employed during the enzyme inhibitory assays affect the final outcomes obtained [23]. Therefore, in order to avoid the potential bias associated with the experimental conditions, it may be more appropriate in the future to build the model with inhibition constant (Ki) values for DPP-IV, which do not depend on the assay conditions employed.

An earlier study based on molecular docking of peptides to the active site of DPP-IV showed that there was a trend ($R^2 = 0.4083$) between the Vina score and the DPP-IV IC_{50} for competitive peptide inhibitors. However, no relationship was found with non-competitive peptide inhibitors of DPP-IV [35]. The importance of taking into account the mode of DPP-IV inhibition of the peptides has again been highlighted by the results obtained herein during QSAR modelling.

The two QSAR models did not allow an accurate prediction of the DPP-IV IC_{50} value as they tended to either over- (v-scale) or underestimate (z-scale) this value. This may be related to the length (≤ 5 amino acids) as well as the limited number of peptides included in the training set. Larger peptides and/or other physicochemical parameters may need to be incorporated in the model. The z5 scale, as developed by Sandberg, Eriksson, Jonsson, Sjöström and Wold [41], which incorporates more physicochemical parameters than the z3 scale used herein, was also used (data not shown). While a better correlation between the predicted and experimentally determined DPP-IV IC_{50} values was obtained ($R^2 = 0.952$) with the z5 scale, the only coefficient of the QSAR model which was significant ($p < 0.05$) was that of the intercept. The ranking of the peptide DPP-IV inhibitory potency was the same when comparing predicted and experimental DPP-IV IC_{50} values (Table 4 and Supplementary Table S1). In addition, both models (z- and v-scale) showed that peptides possessing an hydrophobic N terminal amino acid (i.e., W, I, F and

L) were predicted to be relatively potent inhibitors of DPP-IV, which is in agreement with earlier studies [28, 35, 46]. Inclusion of peptides possessing a wider range of N-terminal amino acids may be advisable in order to verify this hypothesis in future studies.

Several DPP-IV inhibitory peptides with a W at the N-terminal position have been reported in previous studies. A peptide alignment strategy described in an earlier study [28] has been applied to DPP-IV inhibitory peptides with IC_{50} values < 200 μM taking into account more recent information available in the scientific literature (data not shown). Following this sequence alignment, it was seen that an I residue in position N_1 and a P in the N_2 position of the peptides were frequently observed. Overall, there is a good agreement between the prediction obtained with the two QSAR models developed herein and the peptide alignment approaches when applied to DPP-IV inhibitory peptides [28, 46].

Of the peptides used in the training set, a novel and relatively potent DPP-IV inhibitory peptide, IPM (lactoferrin - LF (f 127-129) or LF (f 469-471)), having an IC_{50} value of 69.5 ± 8.7 μM has been reported herein. This peptide has similar structural features (i.e., a tripeptide possessing an IP- N-terminus) as previously identified DPP-IV inhibitory peptides, IPI and IPA, having IC_{50} values of 3.2 [50] and 49 μM [47], respectively. In addition, another relatively potent DPP-IV inhibitory peptide, LPVPQ (β-CN (f 171-175)) originally reported within the milk protein-derived peptides identified in the jejunum of humans [1], having an IC_{50} value of 43.8 ± 8.8 μM was reported herein.

It has been suggested that the insulinotropic properties of milk proteins and milk protein hydrolysates are mediated by an increase in branched chain amino acid (BCAA) and/or in peptide levels in the serum [30, 32]. Other mechanisms, including DPP-IV inhibition induced by milk peptides, may also be involved *in vivo*. Ten selected milk protein-derived peptides identified in the jejunum of humans [1] were shown herein to inhibit DPP-IV *in vitro*. Insulinotropic properties have been reported in small animals following the administration of milk protein-derived peptides and hydrolysates with *in vitro* DPP-IV inhibitory properties [48, 49]. However, to our knowledge, the inhibition of DPP-IV mediated by milk protein-derived peptides has not to date been reported in humans. DPP-IV is present in the gastrointestinal tract, therefore, it is likely that the peptides herein may have an effect on the inhibition of DPP-IV in the gut. Additional human intervention studies are required to verify this assumption.

5 Conclusions

The QSAR approach developed within this study has allowed the ranking of competitive peptides for their potential DPP-IV inhibitory activity. The models were used as a tool to predict the DPP-IV inhibitory properties of peptides which have previously been identified in the intestinal tract of humans following the ingestion of milk/dairy products. The potent DPP-IV inhibitory peptides reported herein may be relevant to human health. It is likely that other peptides identified in the gastrointestinal tract of humans following milk/dairy products ingestion and which were not incorporated in this study may also have the ability to inhibit DPP-IV. While applications of QSAR models in the current context have their limitations (with respect to peptide data set, size and diversity), it is still evident that this approach has shown significant promise in the discovery of new DPP-IV inhibitory peptide sequences. The impact of the milk protein-derived peptides identified in the gut needs to be assessed *in vivo* in order to better understand their role as serum glucose regulating agents.

Abbreviations

ACE, angiotensin converting enzyme; ANN, artificial neural network; BAP: bioactive peptide; BCAA, branched chain amino acid; β-Lg, β-lactoglobulin; BSA, bovine serum albumin; CN, casein; DPP-IV, dipeptidyl peptidase IV; GIP, insulinotropic polypeptide; GLP-1: glucagon-like peptide-1; HCl, hydrochloric acid; IC_{50}, half maximal inhibitory concentration; LF, lactoferrin; MLR, multiple linear regression; PCA, principal component; PLSR, partial least square regression; pNA, p-nitroanilide; PYY, peptide YY; QSAR, quantitative structure activity relationship; SD, standard deviation; T2D, type 2 diabetes; TFA, trifluoroacetic acid; TRIS, tris(hydroxymethyl)aminomethane.

Acknowledgements

The work described herein was supported by Enterprise Ireland under Grant Number TC2013-0001.

Author contributions

A.B. Nongonierma performed all experiments described in this manuscript. A.B. Nongonierma, and R.J. FitzGerald were responsible for experimental design and interpretation of results. A.B. Nongonierma was responsible for writing the manuscript which was critically revised by R.J. FitzGerald. All authors have approved the final article.

Conflicts of interests

The authors declare that they have no conflict of interest.

Supplementary information

Supplementary Table S1. Peptide sequences possessing features of previously identified dipeptidyl peptidase IV (DPP-IV) inhibitors (i.e., with a P at the N_2 position), which have been found in the intestinal tract of humans following the ingestion of milk proteins [1, 4] together with their predicted DPP-IV half maximal inhibitory concentration (IC_{50}). The sequences reported in this Table contain the additional 46 peptides, which were not listed in Table 4.

References

[1] Boutrou R, Gaudichon C, Dupont D, Jardin J, Airinei G, Marsset-Baglieri A, et al. Sequential release of milk protein–derived bioactive peptides in the jejunum in healthy humans. Am J Clin Nutr. 2013;97:1314-23.

[2] Boutrou R, Henry G, Sanchez-Rivera L. On the trail of milk bioactive peptides in human and animal intestinal tracts during digestion: A review. Dairy Sci Technol. 2015;95:815-29.

[3] Chabance B, Jollès P, Izquierdo C, Mazoyer E, Francoual C, Drouet L, et al. Characterization of an antithrombotic peptide from α-casein in newborn plasma after milk ingestion. Br J Nutr. 1995;73:583-90.

[4] Chabance B, Marteau P, Rambaud J, Migliore-Samour D, Boynard M, Perrotin P, et al. Casein peptide release and passage to the blood in humans during digestion of milk or yogurt. Biochimie. 1998;80:155-65.

[5] Chanput W, Nakai S, Theerakulkait C. Introduction of new computer softwares for classification and prediction purposes of bioactive peptides: case study in antioxidative tripeptides. Int J Food Prop. 2010;13:947-59.

[6] Drucker DJ. Enhancing the action of incretin hormones: A new whey forward? Endocrinology. 2006;147:3171-2.

[7] Foltz M, Meynen EE, Bianco V, van Platerink C, Koning TMMG, Kloek J. Angiotensin converting enzyme inhibitory peptides from a lactotripeptide-enriched milk beverage are absorbed intact into the circulation. J Nutr. 2007;137:953-8.
[8] Hatanaka T, Inoue Y, Arima J, Kumagai Y, Usuki H, Kawakami K, et al. Production of dipeptidyl peptidase IV inhibitory peptides from defatted rice bran. Food Chem. 2012;134:797-802.
[9] Hellberg S, Sjoestroem M, Skagerberg B, Wold S. Peptide quantitative structure-activity relationships, a multivariate approach. J Med Chem. 1987;30:1126-35.
[10] Iwaniak A, Minkiewicz P, Darewicz M, Protasiewicz M, Mogut D. Chemometrics and cheminformatics in the analysis of biologically active peptides from food sources. J Funct Foods. 2015;16:334-51.
[11] Jahangiri R, Soltani S, Barzegar A. A review of QSAR studies to predict activity of ACE peptide inhibitors. Pharm Sci. 2014;20:122.
[12] Jao C-L, Hung C-C, Tung Y-S, Lin P-Y, Chen M-C, Hsu K-C. The development of bioactive peptides from dietary proteins as a dipeptidyl peptidase IV inhibitor for the management of type 2 diabetes. BioMedicine. 2015;5:9-15.
[13] Jenssen H. Descriptors for antimicrobial peptides. Expert Opin Drug Discov. 2011;6:171-84.
[14] Jing P, Qian B, He Y, Zhao X, Zhang J, Zhao D, et al. Screening milk-derived antihypertensive peptides using quantitative structure activity relationship (QSAR) modelling and in vitro/in vivo studies on their bioactivity. Int Dairy J. 2014;35:95-101.
[15] Kim H-O, Li-Chan ECY. Quantitative structure−activity relationship study of bitter peptides. J Agric Food Chem. 2006;54:10102-11.
[16] Kost NV, Sokolov OY, Kurasova OB, Dmitriev AD, Tarakanova JN, Gabaeva MV, et al. β-Casomorphins-7 in infants on different type of feeding and different levels of psychomotor development. Peptides. 2009;30:1854-60.
[17] Lacroix IME, Li-Chan ECY. Dipeptidyl peptidase-IV inhibitory activity of dairy protein hydrolysates. Int Dairy J. 2012;25:97-102.
[18] Ledoux N, Mahe S, Dubarry M, Bourras M, Benamouzig R, Tome D. Intraluminal immunoreactive caseinomacropeptide after milk protein ingestion in humans. Food/Nahrung. 1999;43:196-200.
[19] Li Y-W, Li B. Characterization of structure–antioxidant activity relationship of peptides in free radical systems using QSAR models: Key sequence positions and their amino acid properties. J Theor Biol. 2013;318:29-43.
[20] Li Y-W, Li B, He J, Qian P. Structure–activity relationship study of antioxidative peptides by QSAR modeling: the amino acid next to C-terminus affects the activity. J Pept Sci. 2011;17:454-62.
[21] Lin Z-h, Long H-x, Bo Z, Wang Y-q, Wu Y-z. New descriptors of amino acids and their application to peptide QSAR study. Peptides. 2008;29:1798-805.
[22] Meisel H, Fairweather-Tait S, FitzGerald RJ, Hartmann R, Lane CN, McDonagh D, et al. Detection of caseinophosphopeptides in the distal ileostomy fluid of human subjects. Br J Nutr. 2003;89:351-8.
[23] Murray BA, Walsh DJ, FitzGerald RJ. Modification of the furanacryloyl-L-phenylalanylglycylglycine assay for determination of angiotensin-I-converting enzyme inhibitory activity. J Biochem Biophys Methods. 2004;59:127-37.

355 [24] Nakai S, Chan JC, Li-Chan EC, Dou J, Ogawa M. Homology similarity analysis of
356 sequences of lactoferricin and its derivatives. J Agric Food Chem. 2003;51:1215-23.
357 [25] Nongonierma AB, FitzGerald RJ. Dipeptidyl peptidase IV inhibitory and antioxidative
358 properties of milk-derived dipeptides and hydrolysates. Peptides. 2013;39:157-63.
359 [26] Nongonierma AB, FitzGerald RJ. Inhibition of dipeptidyl peptidase IV (DPP-IV) by proline
360 containing peptides. J Funct Foods. 2013;5:1909–17.
361 [27] Nongonierma AB, FitzGerald RJ. Inhibition of dipeptidyl peptidase IV (DPP-IV) by
362 tryptophan containing dipeptides. Food Funct. 2013;4:1843-9.
363 [28] Nongonierma AB, FitzGerald RJ. An *in silico* model to predict the potential of dietary
364 proteins as sources of dipeptidyl peptidase IV (DPP-IV) inhibitory peptides. Food Chem.
365 2014;165:489–98.
366 [29] Nongonierma AB, FitzGerald RJ. Susceptibility of milk protein-derived peptides to
367 dipeptidyl peptidase IV (DPP-IV) hydrolysis. Food Chem. 2014;145:845–52.
368 [30] Nongonierma AB, FitzGerald RJ. Bioactive properties of milk proteins in humans: A
369 review. Peptides. 2015;73:20-34.
370 [31] Nongonierma AB, FitzGerald RJ. Milk proteins as a source of tryptophan-containing
371 bioactive peptides. Food Funct. 2015;6:2115-27.
372 [32] Nongonierma AB, FitzGerald RJ. The scientific evidence for the role of milk protein-
373 derived bioactive peptides in humans: A Review. J Funct Foods. 2015;640:640–56.
374 [33] Nongonierma AB, FitzGerald RJ. Prospects for the management of type 2 diabetes using
375 food protein-derived peptides with dipeptidyl peptidase IV (DPP-IV) inhibitory activity. Curr
376 Opin Food Sci. 2016;8:19-24.
377 [34] Nongonierma AB, FitzGerald RJ. Strategies for the discovery, identification and validation
378 of milk protein-derived bioactive peptides. Trends Food Sci Technol. 2016;50:26-43.
379 [35] Nongonierma AB, Mooney C, Shields DC, FitzGerald RJ. *In silico* approaches to predict the
380 potential of milk protein-derived peptides as dipeptidyl peptidase IV (DPP-IV) inhibitors.
381 Peptides. 2014;57:43-51.
382 [36] Panchaud A, Affolter M, Kussmann M. Mass spectrometry for nutritional peptidomics:
383 How to analyze food bioactives and their health effects. J Proteomics. 2012;75:3546-59.
384 [37] Pripp AH, Ardö Y. Modelling relationship between angiotensin-(I)-converting enzyme
385 inhibition and the bitter taste of peptides. Food Chem. 2007;102:880-8.
386 [38] Pripp AH, Isaksson T, Stepaniak L, Sørhaug T. Quantitative structure-activity relationship
387 modelling of ACE-inhibitory peptides derived from milk proteins. Eur Food Res Technol.
388 2004;219:579-83.
389 [39] Pripp AH, Isaksson T, Stepaniak L, Sørhaug T, Ardö Y. Quantitative structure activity
390 relationship modelling of peptides and proteins as a tool in food science. Trends Food Sci
391 Technol. 2005;16:484-94.
392 [40] Rekdal Ø, Andersen J, Vorland LH, Svendsen JS. Construction and synthesis of
393 lactoferricin derivatives with enhanced antibacterial activity. J Pept Sci. 1999;5:32-45.
394 [41] Sandberg M, Eriksson L, Jonsson J, Sjöström M, Wold S. New chemical descriptors
395 relevant for the design of biologically active peptides. A multivariate characterization of 87
396 amino acids. J Med Chem. 1998;41:2481-91.
397 [42] Silveira ST, Martínez-Maqueda D, Recio I, Hernández-Ledesma B. Dipeptidyl peptidase-IV
398 inhibitory peptides generated by tryptic hydrolysis of a whey protein concentrate rich in β-
399 lactoglobulin. Food Chem. 2013;141:1072–7.

[43] Strøm MB, Stensen W, Svendsen JS, Rekdal Ø. Increased antibacterial activity of 15-residue murine lactoferricin derivatives. J Pept Res. 2001;57:127-39.
[44] Svedberg J, de Haas J, Leimenstoll G, Paul F, Teschemacher H. Demonstration of β-casomorphin immunoreactive materials in in vitro digests of bovine milk and in small intestine contents after bovine milk ingestion in adult humans. Peptides. 1985;6:825-30.
[45] Tian M, Fang B, Jiang L, Guo H, Cui J, Ren F. Structure-activity relationship of a series of antioxidant tripeptides derived from β-Lactoglobulin using QSAR modeling. Dairy Sci Technol. 2015;95:451-63.
[46] Tulipano G, Faggi L, Nardone A, Cocchi D, Caroli AM. Characterisation of the potential of β-lactoglobulin and α-lactalbumin as sources of bioactive peptides affecting incretin function: *In silico* and *in vitro* comparative studies. Int Dairy J. 2015;48:62-72.
[47] Tulipano G, Sibilia V, Caroli AM, Cocchi D. Whey proteins as source of dipeptidyl dipeptidase IV (dipeptidyl peptidase-4) inhibitors. Peptides. 2011;32:835-8.
[48] Uchida M, Ohshiba Y, Mogami O. Novel dipeptidyl peptidase-4–inhibiting peptide derived from β-lactoglobulin. J Pharmacol Sci. 2011;117:63-6.
[49] Uenishi H, Kabuki T, Seto Y, Serizawa A, Nakajima H. Isolation and identification of casein-derived dipeptidyl-peptidase 4 (DPP-4)-inhibitory peptide LPQNIPPL from gouda-type cheese and its effect on plasma glucose in rats. Int Dairy J. 2012;22:24-30.
[50] Umezawa H, Aoyagi T, Ogawa K, Naganawa H, Hamada M, Takeuchi T. Diprotins A and B, inhibitors of dipeptidyl aminopeptidase IV, produced by bacteria. J Antibiot. 1984;37:422-5.
[51] Wu J, Aluko RE. Quantitative structure-activity relationship study of bitter di- and tri-peptides including relationship with angiotensin I-converting enzyme inhibitory activity. J Pept Sci. 2007;13:63-9.
[52] Wu J, Aluko RE, Nakai S. Structural requirements of angiotensin I-converting enzyme inhibitory peptides: Quantitative structure-activity relationship study of di-and tripeptides. J Agric Food Chem. 2006;54:732-8.
[53] Wu J, Aluko RE, Nakai S. Structural requirements of angiotensin I-converting enzyme inhibitory peptides: quantitative structure-activity relationship modeling of peptides containing 4-10 amino acid residues. QSAR Comb Sci. 2006;25:873-80.

Table captions

Table 1. Dipeptidyl peptidase IV (DPP-IV) half maximal inhibitory concentration (IC_{50}) and mode of inhibition of milk protein-derived peptides (this information was used to build the quantitative structure activity relationship (QSAR) models).

Table 2. Statistical significance of the partial least square regression (PLSR) analysis linking the half maximal inhibitory concentration (IC_{50}) for dipeptidyl peptidase IV (DPP-IV) of the competitive peptide inhibitors used to generate the quantitative structure activity relationship (QSAR) models.

Table 3. Coefficients of the partial least square regression (PLSR) analysis used to generate the quantitative structure activity relationship (QSAR) models linking the half maximal inhibitory concentration (IC_{50}) for dipeptidyl peptidase IV (DPP-IV) with the peptide descriptors.

Table 4. β-Casein derived peptide sequences which have previously been identified in the intestinal tract of humans following the ingestion of milk proteins [1] together with their predicted and experimentally determined dipeptidyl peptidase IV (DPP-IV) half maximal inhibitory concentration (IC_{50}). This table contains the data obtained with the ten peptides selected for the confirmatory study.

Table 1

Peptide sequence[1]	Milk protein fragment[2]	DPP IV IC_{50} (μM)[3]	Mode of inhibition	Peptide length	Reference
IPI	κ-CN (f 26-28)	3.5	competitive	3	[29]
IPIQY	κ-CN (f 26-30)	35.2	competitive	5	[29]
FLQP	β-CN (f 102-105)	65.3	competitive	4	[26]
IPM	LF (f 127-129), LF (f 469-471)	69.5 ± 8.7	competitive	3	this study
LPYPY	κ-CN (f 56-60)	108.3	competitive	5	[29]
HL	diverse	143.2	competitive	2	[25]
IP	diverse	149.6	competitive	2	[29]
VA	diverse	168.2	competitive	2	[25]
YPYY	κ-CN (f 58-61)	194.4	competitive	4	[29]
LPL	β-CN (f 150-152), β-CN (f 152-154)	241.4	competitive	3	[29]
YPY	κ-CN (f 58-60)	243.7	competitive	3	[29]
WY	α_{s1}-CN (f 164-165), β-Lg (f 19-20)	281.0	competitive	2	[27]
LPLPL	β-CN (f 150-154)	325.0	competitive	5	[29]
VP	diverse	380.3 ± 28.4	competitive	2	this study
FL	diverse	399.6	competitive	2	[25]
IPSK	LF (f 310-313)	406.8 ± 50.4	competitive	4	this study
IPPL	β-CN (f 89-92)	428.9 ± 36.1	uncompetitive	4	this study
LPP	β-CN (f 166-168), BSA (f 301-303)	563.3 ± 50.4	non-competitive	3	this study
VLGP	β-CN (f 212-215)	580.4	competitive	4	[26]
RP	diverse	657.2 ± 38.4	competitive	2	this study
YP	diverse	658.1	competitive	2	[29]
FP	diverse	682.5 ± 79.4	mixed	2	this study
LP	diverse	712.5	competitive	2	[29]
AL	diverse	882.1	competitive	2	[25]
HP	diverse	902.8 ± 93.8	mixed	2	this study
LW	α_{s1}-CN (f 198-199)	993.4	competitive	2	[27]
LQP	β-CN (f 103-105)	1181.1	competitive	3	[26]
SL	diverse	2517.1	competitive	2	[25]
GL	diverse	2615.0	competitive	2	[25]

EK	diverse	3216.8	competitive	2	[25]

[1]DPP-IV inhibitory peptide sequences abbreviated using the one letter amino acid code. Peptides are ordered by increasing IC_{50} value. The competitive peptide inhibitors were used to generate the quantitative structure activity relationship (QSAR) models.
[2]BSA: bovine serum albumin, CN: casein, β-Lg: β-lactoglobulin, LF: lactoferrin.
[3]All DPP-IV IC_{50} values listed were obtained using the same experimental protocol as described in Nongonierma and FitzGerald [25].

Table 2

Parameter	z-scale	v-scale
n	26	26
R^2 model	0.829	0.815
p model	0.00279	0.00437
Root mean square error	484	504
F	5.27	4.78
R^2 cross validation	0.775	0.754

Table 3

Amino acid descriptors	Coefficient	Model coefficient	Standard error	tStat	p^1
z-scale [9]	c	1073	196	5.48	0.0001
	$\alpha_{1,1}$	146	60	2.42	0.03
	$\alpha_{1,2}$	-359	208	-1.73	0.11
	$\alpha_{1,3}$	37	210	0.18	0.86
	$\alpha_{2,1}$	213	96	2.22	0.04
	$\alpha_{2,2}$	-247	434	-0.57	0.58
	$\alpha_{2,3}$	-233	138	-1.69	0.12
	$\beta_{1,1}$	-4	64	-0.06	0.95
	$\beta_{1,2}$	155	200	0.77	0.45
	$\beta_{1,3}$	-41	115	-0.36	0.72
	$\beta_{2,1}$	-178	107	-1.66	0.12
	$\beta_{2,2}$	545	413	1.32	0.21
	$\beta_{2,3}$	-129	132	-0.98	0.35
v-scale [21]	c	1341	489	2.74	0.02
	$\alpha_{1,1}$	-318	2455	-0.13	0.90
	$\alpha_{1,2}$	7053	4614	1.53	0.15
	$\alpha_{1,3}$	-369	131	-2.82	0.01
	$\alpha_{2,1}$	3527	1859	1.90	0.08
	$\alpha_{2,2}$	518	2353	0.22	0.83
	$\alpha_{2,3}$	-478	254	-1.88	0.08
	$\beta_{1,1}$	-1636	2498	-0.65	0.52
	$\beta_{1,2}$	-2414	3337	-0.72	0.48
	$\beta_{1,3}$	56	144	0.38	0.71
	$\beta_{2,1}$	-396	1642	-0.24	0.81
	$\beta_{2,2}$	-694	2232	-0.31	0.76
	$\beta_{2,3}$	73	260	0.28	0.78

[1]Coefficients with a $p < 0.05$ are significantly different from 0.

Table 4

Peptide Sequence[1]	β-Casein fragment	DPP IV IC_{50} (μM)			Mode of DPP-IV inhibition[4]
		Predicted1[2]	Predicted2[2]	Experimental[3]	
VPGEIVE	8-14	-109	844	224.5 ± 66.3 c	competitive
YPFPGP	60-65	nd	nd	749.2 ± 122.3 f	mixed-type
LPQNIPPLT	70-78	-195	519	205.2 ± 32.5 c	competitive
IPPLTQT	74-80	nd	nd	465.1 ± 73.7 e	uncompetitive
TPVVVPP	80-86	nd	nd	1408.9 ± 179.0 h	mixed-type
YPVEPF	114-119	nd	nd	124.7 ± 9.1 b	mixed-type
LPLPLL	135-140	6	556	371.5 ± 60.6 d	competitive
QPHQPLPPT	146-154	nd	nd	1754.8 ± 375.0 i	mixed-type
QPLPPT	149-154	447	1198	1013.8 ± 162.4 g	competitive
LPVPQ	171-175	-491	97	43.8 ± 8.8 a	competitive

nd: not determined

[1]Peptide sequences abbreviated using the one letter amino acid code.

[2]DPP-IV IC_{50} value predicted with the QSAR model developed with the z-scale (predicted 1) and v-scale (predicted 2) amino acid descriptors as described in Table 3. The predicted values were only determined for the competitive DPP-IV inhibitory peptides. The IC_{50} of the positive control IPI was of 3.73 ± 0.56 μM.

[3]Mean ± SD (n=3). Figures with different letters are significantly different at $p < 0.05$.

[4]The mode of inhibition was determined with the Lineweaver and Burk double reciprocal representation.

Figure captions

Figure 1. Lineweaver and Burk double reciprocal plots for dipeptidyl peptidase IV (DPP-IV) inhibition with (A) IPPL (uncompetitive), (B) HP (mixed-type), (C) LPP (non-competitive) and (D) RP (competitive) evaluated at concentrations corresponding to their IC_{50} divided by 8 and 40. Each point is the mean ± SD (n=4).

Figure 2. Quantitative structure activity relationship (QSAR) models generated with the competitive dipeptidyl peptidase IV (DPP-IV) inhibitors described in Table 1, using the amino acid descriptors developed by (A) Hellberg, Sjoestroem, Skagerberg and Wold [9] (z-scale) and (B) Lin, Long, Bo, Wang and Wu [21] (v-scale).

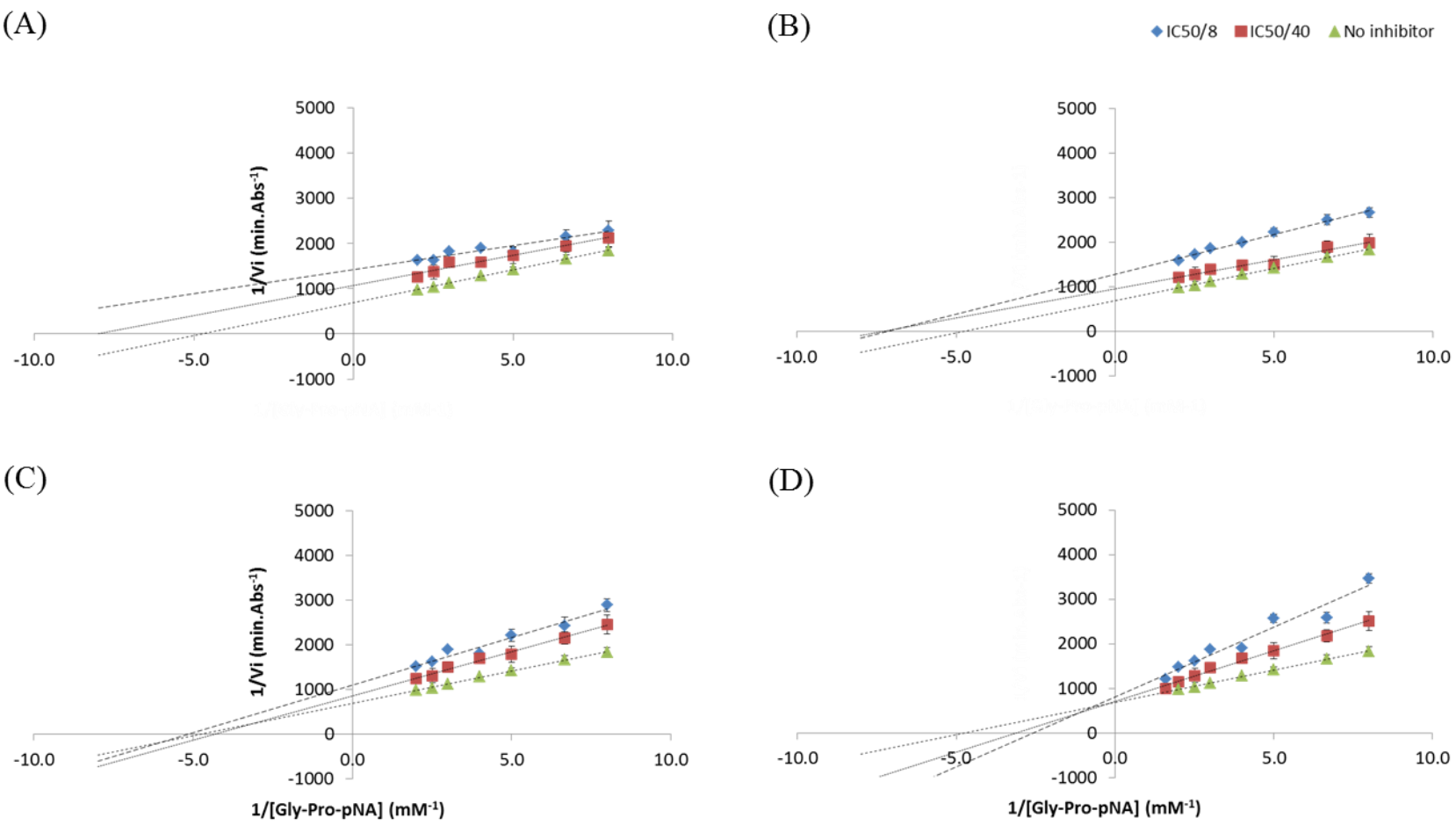
(A)
(B)
(C)
(D)
IC50/8
IC50/40
No inhibitor
1/Vi (min.Abs-1)
1/[Gly-Pro-pNA] (mM-1)
5000
4000
3000
2000
1000
0
-1000
-10.0
-5.0
0.0
5.0
10.0

Figure 1

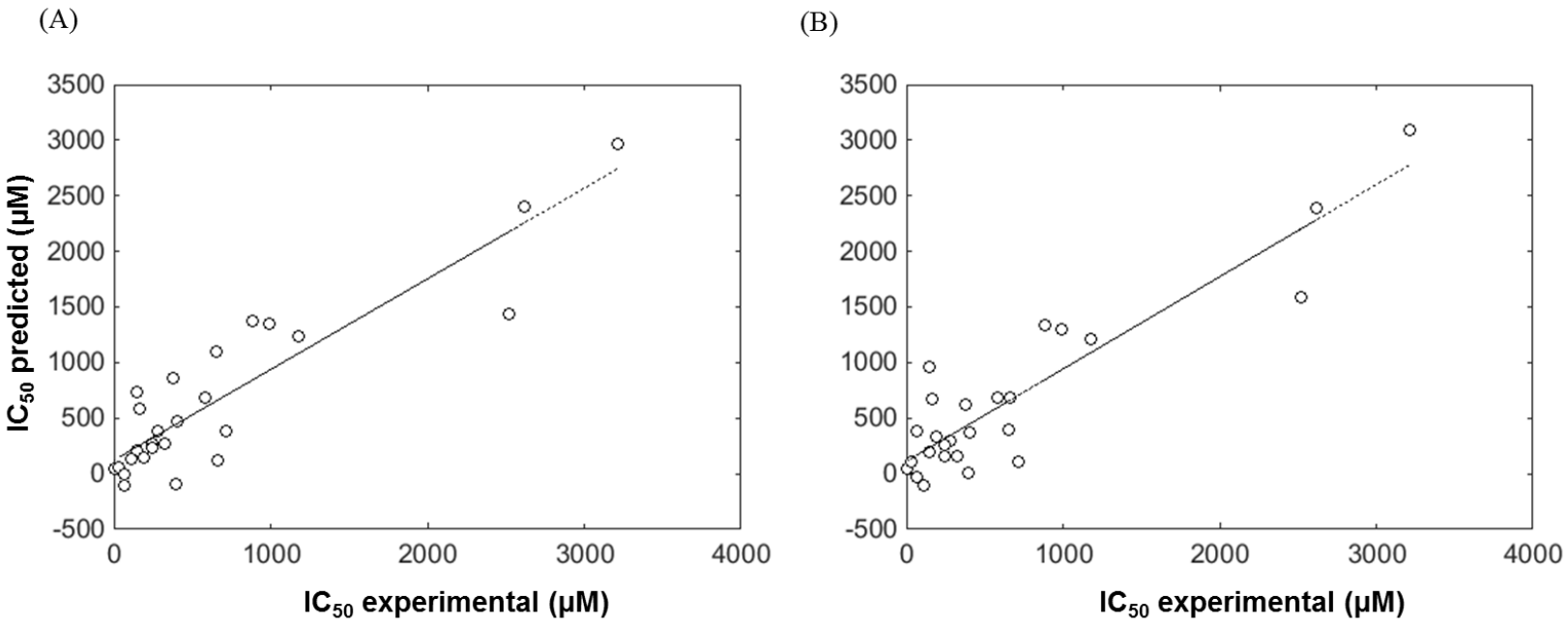
(A)
(B)
3500
3000
2500
2000
1500
1000
500
0
-500
0
1000
2000
3000
4000
IC50 predicted (μM)
IC50 experimental (μM)
IC50 experimental (μM)

Figure 2

Supplementary Table S1. Peptide sequences possessing features of previously identified dipeptidyl peptidase IV (DPP-IV) inhibitors (i.e., with a P at the N_2 position), which have been found in the intestinal tract of humans following the ingestion of milk proteins [1, 4] together with their predicted DPP-IV half maximal inhibitory concentration (IC_{50}). The sequences reported in this Table contain the additional 46 peptides, which were not listed in Table 4.

Peptide Sequence[1]	**Parent protein**[2]	**Fragment**	**DPP IV IC_{50} (μM)**			
			Predicted1[3]	**Predicted2**[3]	**Experimental**[4]	**Reference**[4]
YPFPGPI	β-CN	60-66	217	407		
YPFPGPIP	β-CN	60-67	371	572		
YPFPGPIPN	β-CN	60-68	-978	237	670	[49]
FPGPIP	β-CN	62-67	-883	271		
FPGPIPN	β-CN	62-68	-2232	-63	260	[49]
LPQNIPPL	β-CN	70-77	156	274	46	[49]
LPQNIPPLTQT	β-CN	70-80	-32	428		
LPQNIPPLTQTPV	β-CN	70-82	-888	271		
IPPLTQTPV	β-CN	74-82	-693	8	1300	[49]
IPPLTQTPVVVPPFLQPE	β-CN	74-91	-285	-170		
IPPLTQTPVVVPPFLQPEV	β-CN	74-92	-691	303		
PPLTQ	β-CN	75-79	-496	2556		
PPLTQTPV	β-CN	75-82	-1020	2139		
PPLTQTPVVVPPFLQPE	β-CN	75-91	-613	1961		
PPLTQTPVVVPPFLQPEV	β-CN	75-92	-1018	2434		
TPVVVPPF	β-CN	80-87	1630	670		
TPVVVPPFLQP	β-CN	80-90	527	743		
TPVVVPPFLQPE	β-CN	80-91	-220	538		
TPVVVPPFLQPEV	β-CN	80-92	-626	1012		
TPVVVPPFLQPEVM	β-CN	80-93	79	1135		
VPPFLQPE	β-CN	84-91	266	413		
VPPFLQPEV	β-CN	84-92	-140	886		
PPFLQ	β-CN	85-89	-773	2246		
PPFLQPE	β-CN	85-91	-613	1961		

PPFLQPEV	β-CN	85-92	-1018	2434		
PPFLQPEVM	β-CN	85-93	-313	2557		
MPFPK	β-CN	109-113	511	-112		
YPVEPFT	β-CN	114-120	473	376		
YPVEPFTESQ	β-CN	114-123	-458	930		
LPLPL	β-CN	135-139	156	274	325	[29]
QPHQPLPPTVMFPPQS	β-CN	146-161	-958	1399		
QPHQPLPPTVMFPPQSV	β-CN	146-162	-727	1963		
QPLPPTV	β-CN	149-155	-267	1823		
QPLPPTVMFPPQ	β-CN	149-160	2	1057		
EPVLGPV	β-CN	195-201	-301	1149		
EPVLGPVRGPF	β-CN	195-205	1956	1102		
EPVLGPVRGPFP	β-CN	195-206	994	1021		
EPVLGPVRGPFPI	β-CN	195-207	439	1184		
EPVLGPVRGPFPII	β-CN	195-208	192	1549		
VPPFLQPEV	β-CN	84-92	-140	886		
YPVEPF	β-CN	114-119	1735	325		
FPEVFGKE	α_{s1}-CN	28-35	-1082	-70		
VPQLEIVPN	α_{s1}-CN	106-114	-596	456		
EPMIGV	α_{s1}-CN	133-138	-1200	1609		
YPFPGPI	β-CN	60-66	217	407		
HPIKHQGLPQEV	α_{s1}-CN	4-15	-838	917		

[1]Peptide sequences abbreviated using the one letter amino acid code.

[2]CN: casein.

[3]DPP-IV IC_{50} value predicted with the QSAR model developed with the z-scale (predicted 1) and v-scale (predicted 2) amino acid descriptors as described in Table 3.

[4]Experimental DPP-IV IC_{50} values as reported in previous studies.